Europ. J. Combinatorics (1996) **17,** 255–264

Minimum Spanning Trees and Types of Dissimilarities

BRUNO LECLERC

This paper is devoted to structural relations between types of dissimilarities and the corresponding minimum spanning trees (MSTs). It is first shown that dissimilarity preorders inducing directly hierarchical classification or seriation may be characterized in terms of MSTs. Several MST-preserving mappings (most of them being anticlosures) onto dissimilarities of special types are constructed. Some aspects of the role of MSTs in lattices of ultrametrics are presented, including their use in a proof of the semimodularity of these lattices.

On s'intéresse aux relations structurelles entre certains types de dissimilarités et les arbres minimums correspondants. On montre d'abord que les préordonnances qui induisent directement une classification hiérarchique ou une sériation se caractérisent en termes d'arbres minimums. On construit des applications (la plupart étant des ouvertures) préservant les arbres minimums et dont les images correspondent à des types particuliers de dissimilarités. Quelques aspects du rôle des arbres minimums dans les treillis d'ultramétriques sont présentés, incluant leur usage dans une démonstration de la semimodularité de ces treillis.

1. INTRODUCTION

According to the historical survey of Graham and Hell [12], minimum spanning trees (abbreviated as MSTs in the sequel) were discovered in 1926 by Boruvka [4] in the study of the minimum length connected subgraph problem. After several rediscoverings of Boruvka's solution, the uses of MSTs in other problems relevant to Operational Research were pointed out, together with the strength of their mathematical properties. Among these properties are the purely ordinal nature of MSTs, and their relation with the greedy algorithm through the fact that trees are a special case of matroid bases [10, 25].

Besides Operational Research, Florek *et al.* [9] opened in the early 1950s a second domain of applications for MSTs, the analysis of data consisting of a dissimilarity index on a finite set. On the one hand, MSTs are an important tool in many problems of combinatorial optimization issued from classification (for recent and typical examples, see [13–15]. On the other hand, there are structural relations between some types of dissimilarities and their MSTs, the first observation of this fact being in a celebrated paper of Gower and Ross [11] connecting MSTs with a classical clustering algorithm and, in fact, with ultrametrics. This paper is devoted to this last aspect; some parts of its contents have been previously published, sometimes without the proofs, in Leclerc [16–19] or are to appear [22].

Basic definitions about dissimilarities and dissimilarity preorders on a finite set X are recalled in Section 2.1, and a definition of MSTs, in the context of a dissimilarity preorder R, in Section 2.2. We point out in Section 3 that MSTs exist when the preorder R, possibly not complete, is issued from a hierarchical classification or from a linear order on X. The corresponding dissimilarities are, respectively, the ultrametrics and the Robinson dissimilarities. It is interesting to have transformations that map a given dissimilarity into another one of a prescribed type with preservation of MSTs, because any MST provides a hierarchical classification (by the single linkage algorithm) and partial seriations (as seen in Section 3.2). This is done in Section 4, with some emphasis on the definition of lattice structures; sets of dissimilarities with a given MST

0195-6698/96/020255 + 10 $18.00/0

are also considered. The last section is devoted to the special case of the lattice of ultrametrics, for the study of which MSTs constitute a particularly efficient tool.

2. Definitions

2.1. *Dissimilarities and dissimilarity preorders.* Throughout the paper, X will be a finite set with n elements, and $X^{(2)}$ the set of all the unordered pairs of elements of X. A dissimilarity on X is a function $d: X^2 \to \mathbb{R}_+$ (the set of the real non-negative numbers) satisfying the properties $d(x, x) = 0$ for all $x \in X$ and $d(x, y) = d(y, x)$ for all $x, y \in X$. Here, dissimilarities will be considered in the equivalent form of functions from $X^{(2)}$ into $\mathbb{R}_+$.

A *dissimilarity preorder* (abbreviated to DP in the sequel) R on X is a preorder (reflexive and transitive binary relation) on $X^{(2)}$. Here, for two pairs $xy, x'y' \in X^{(2)}$, $(xy, x'y') \in R$ (denoted $xyRx'y'$) means that x and y are more or equally similar to each other than x' and y' are. The relation R is not assumed to be complete: it may exist *incomparable pairs* e, e' with neither eRe' nor $e'Re$. The symmetric and the asymmetric parts of R are respectively denoted as R^s and R^a: eR^se' (resp. eR^ae') implies eRe' and $e'Re$ (resp. not $e'Re$). A pair e is *minimum* for R if eRe' for all pairs $e' \in X^{(2)}$, and *minimal* if, for any $e' \in X^{(2)}$, $e'Re$ implies eRe'.

A dissimilarity d naturally induces a complete DP R_d *by*: eR_de' iff $d(e) \leqslant d(e')$. More generally, a preorder R on $X^{(2)}$ and a dissimilarity d on X are said to be *compatible* if eRe' implies $d(e) \leqslant d(e')$ and eR^se' implies $d(e) = d(e')$.

A dissimilarity is *even*, a *q-semimetric* ($q \in \mathbb{R}_+$, $q > 0$), an *ultrametric* or a *tree dissimilarity* if it satisfies, respectively, the following properties (E), (qM), (U) or (T):

(E) For all distinct $x, y \in X$, $d(xy) = 0$ implies: for all $z \in X$, $d(xz) = d(yz)$;

(qM) For all distinct $x, y, z \in X$, $(d(xz))^q \leqslant (d(xy))^q + (d(yz))^q$;

(U) For all distinct $x, y, z \in X$, $d(xz) \leqslant \max(d(xy), d(yz))$;

(T) For all distinct $x, y, z, w \in X$, $d(xy) + d(zw) \leqslant \max\{d(xz) + d(yw), d(xw) + d(yz)\}$.

Since the function $(a^q + b^q)^{1/q}$ of q is decreasing for positive a, b, the implication (qM) $\Rightarrow$ (q'M) holds for $q' \leqslant q$. The *semimetrics* correspond to $q = 1$ (the property (qM) is written (M) in this case), the even dissimilarities to $q = 0$ (such a convention is natural, since (qM) $\Rightarrow$ (E) for all $q > 0$) and the ultrametrics to q infinite. The implication (U) $\Rightarrow$ (T) is easy to prove. If the elements of X are not assumed to be distinct, then (T) $\Rightarrow$ (M) also holds. It is well known that Property (T) means that two among the three sums $d(xy) + d(zw)$, $d(xz) + d(yw)$ and $d(xw) + d(zv)$ are equal and at least equal to the third, while (U) means that two of the numbers $d(xy)$, $d(xz)$ and $d(yz)$ are equal and at least equal to the third.

The sets of all dissimilarities, even dissimilarities, semimetrics, q-semimetrics, tree dissimilarities and ultrametrics on X will be denoted, respectively, as $\mathscr{D}$, $\mathscr{E} = \mathscr{M}_0$, $\mathscr{M} = \mathscr{M}_1$, $\mathscr{M}_q$, $\mathscr{T}$ and $\mathscr{U} = \mathscr{M}_\infty$. These sets are naturally endowed with the *pointwise order*:

$$\text{For all } d, d' \in \mathscr{D},\ d \leqslant d' \Leftrightarrow d(xy) \leqslant d'(xy) \qquad \text{for all distinct } x, y \in X.$$

This order is a distributive lattice, with the pointwise join (or least upper bound, denoted by $\vee$) and meet (or greatest lower bound, denoted $\wedge$) operations:

$$\text{for all distinct } x, y \in X, \qquad (d \vee d')(xy) = \max(d(xy), d'(xy));$$
$$(d \wedge d')(xy) = \min(d(xy), d'(xy)).$$

Recall that a lattice $\mathscr{L}$ is *distributive* if it satisfies the *distributivity laws*: for all d, d', $t \in \mathscr{L}$, $(d \vee d') \wedge t = (d \wedge t) \vee (d' \wedge t)$, or, equivalently, $(d \wedge d') \vee t = (d \vee t) \wedge (d' \vee t)$. A finite lattice $\mathscr{L}$ is *lower semimodular* if, for every d, $d' \in \mathscr{L}$, $d \prec d \vee d'$ and $d' \prec d \vee d'$ imply $d \wedge d' \prec d$ and $d \wedge d' \prec d'$, where $\prec$ denotes the *covering relation* on $\mathscr{L}$: $t \prec d$ means that $t < d$ and $t \leqslant t' < d$ implies that $t = t'$. An extension of lower semimodularity to infinite lattice will be mentioned in Section 5. Distributive lattices are lower semimodular. For other definitions and terminology about lattices, see [3].

2.2. *Minimal and minimum spanning trees.* Consider the complete graph $K_X = (X, X^{(2)})$ on X. A *spanning tree on* X is a subset A of $X^{(2)}$ such that the subgraph (X, A) is a tree-graph (a connected and acyclic graph). The unique chain of A between two distinct elements x, y of X is denoted $A(xy)$. Let $\mathscr{A}$ be the set of all the spanning trees on X. A relation $\bar{R}$ on $\mathscr{A}$ is associated to any DP R by:

for A, $A' \in \mathscr{A}$, $A \,\bar{R}\, A'$

$\Leftrightarrow$ there exists a bijection $\beta: A \to A'$ with $aR\beta(a)$ for all $a \in A$.

We do not prove here the following basic facts: $\bar{R}$ is a preorder, and an order if R is an order. A spanning tree A on X is said to be a *minimum spanning tree* (MST) *for* R if it is minimum for $\bar{R}$, and a *minimal spanning tree* (mst) *for* R if it is minimal for $\bar{R}$. When no confusion is possible, the mention 'for R' will be omitted in the sequel.

An exchange relation $\Delta_A \subseteq A \times (X^{(2)} - A)$ is associated with any spanning tree $A \in \mathscr{A}$. It is defined by: for all $a \in A$, $b \in X^{(2)} - A$, $a\Delta_A b \Leftrightarrow (A - \{a\}) \cup \{b\} \in \mathscr{A} \Leftrightarrow a \in A(b)$; we then set $(A - \{a\}) \cup \{b\} = A_{ab}$. Two facts are useful: (i) for each $b \notin A$, the set $\{b\} \cup \{a \in A: a\Delta_A b\} = \{b\} \cup A(b)$ is a cycle, denoted $C_{A,b}$; (ii) for each $a \in A$, the set $\{a\} \cup \{b \in X^{(2)} - A: a\Delta_A b\}$ is the cocycle, denoted $D_{A,a}$, of all the edges with one extremity in each of the connected components of the graph $(X, A - \{a\})$. The following characterizations of msts and MSTs have been obtained by Flament and Leclerc [8] and Leclerc [19], some of them in the more general context of bases of a matroid defined on a finite set.

PROPOSITION 2.1. *Let A be a spanning tree and let R be a* DP *on X. Then, the following three conditions are equivalent*:
(1) *A is an* MST;
(2) $\Delta_A \subseteq R$;
(3) *for any dissimilarity d on X compatible with R, the quantity $\sum_{a \in A'} d(a)$ is minimized in $\mathscr{A}$ by the spanning tree A.*
The following three conditions are equivalent:
(4) *A is an* mst;
(5) *every directed cycle of the graph $(X^{(2)}, R \cup \Delta_A)$ is a directed cycle of the graph $(X^{(2)}, R^s \cup \Delta_A)$;*
(6) *there exists a dissimilarity d on X compatible with R such that the quantity $\sum_{a \in A'} d(a)$ is minimized in $\mathscr{A}$ by the spanning tree A.*

The set of all the MSTs corresponding to a DP R (or to a dissimilarity d) is denoted as $\mathscr{A}_{M,R}$ (or $\mathscr{A}_{M,d}$). If A is an mst for a complete DP R, then (5) implies (2); in fact, it is a classical result of Matroid Theory [10] that a complete DP R has at least one MST. On the other hand, it follows from the definitions that a dissimilarity order has at most one MST. Such a dissimilarity order P_A is canonically associated to any spanning tree A on X by: for e, $e' \in X^{(2)}$, eP_Ae' iff $A(e) \subseteq A(e')$. Obviously, the above condition (2) is satisfied by this order, which has A as unique MST.

3. Characterizations in Terms of MSTs

3.1. *MSTs and hierarchical classification.* Hierarchies constitute a formalization of classification trees and are extensively considered in the mathematics of classification. A *hierarchy* H on X is a family of subsets (clusters) of X satisfying the following properties: $X \in H$; $\emptyset \notin H$; for all $x \in X$, $\{x\} \in H$; for all $h, h' \in H$, $h \cap h' \in \{h, h', \emptyset\}$. A hierarchy H, endowed with the inclusion order, y is a tree semilattice: for $x, y \in X$, there exists a lowest cluster xHy of H including both x and y. We then associate to H a DP R_H, generally not complete, by:

$$\text{for any } xy, x'y' \in X^{(2)}, \qquad xyR_Hx'y' \Leftrightarrow xHy \subseteq x'Hy'.$$

A DP R on X is said to be *hierarchical* if $R \supseteq R_H$ for some hierarchy H on X.

Theorem 3.1. *For a* DP R *on* X *such that* xxR^axy *for all distinct* $x, y \in X$, *the following four properties are equivalent:*
(1) R *is hierarchical*;
(2) *for all distinct* $x, y, z \in X$, $xzRxy$ *or* $xzRyz$;
(3) *for all distinct* $x_1, x_2, \ldots, x_k \in X$, *there exists* $i \in \{1, \ldots, k-1\}$ *with* $x_1x_kRx_ix_{i+1}$;
(4) *the equality* $\bigcup\{A \in \mathscr{A}_{M,R}\} = X^{(2)}$ *holds.*

Proof. (1) $\Rightarrow$ (2): let H be a hierarchy such that $R_H \subseteq R$; consider three distinct elements x, y, z of X, and $h = xHy$, $h' = yHz$. Since $h \cap h'$ is not empty, one has, say, $h \cap h' = h$, that is $h \subseteq h'$, which implies $xHz \subseteq h'$ and $xzRxy$. The case $h' \subseteq h$ leads to $xzRyz$.

(2) $\Rightarrow$ (3): from (2), (3) is satisfied for $k \leqslant 3$; assume that it is satisfied up to $k-1$. Then, from (2) again, one has $x_1x_k \leqslant x_1x_{k-1}$ or $x_1x_k \leqslant x_{k-1}x_k$. It follows from the induction hypothesis in the second case, and directly in the first one, that (3) is satisfied.

(3) $\Rightarrow$ (4): let A be an mst and a pair $b \notin A$. From (3), there exists $a \in A(b)$ with bRa. One then has $A_{ab}\bar{R}A$ and, since A is an mst, $A\bar{R}A_{ab}$; so, A_{ab} is also an mst and every pair belongs to an mst; moreover, aR^sb. It remains to show that A is in fact an MST. For an element a', distinct from a, of the chain $A(b)$, $a'R^aa$ would imply $a'R^ab$, and $A_{a'b}\bar{R}A$ with not $A\bar{R}A_{a'b}$, a contradiction with the assumption that A is an mst. If a' is not comparable with a, then, there exists a subchain $A(b')$ of $A(b)$ with exactly two incomparable maximal elements, say, a and a'. For the same reasons as above, one has aR^sb' or $a'R^sb'$. In the first case, the chain $(A(b') - \{a'\}) \cup \{b'\}$ does not satisfy (3); the second case is similar. Thus, a is maximum for R in $A(b)$, which implies $a'Rb$ for all a' such that $a'\Delta_Ab$. Finally, A satisfies the property (2) of Proposition 2.1 and is an MST.

(4) $\Rightarrow$ (3): let C be a chain of K_X between two distinct vertices x and y, and A an MST such that $xy \in A$. The cocycle $D_{A,xy}$ intersects the cycle $C \cup \{xy\}$ in xy and in at least one other pair b. Then, $xy\Delta_Ab$ implies $xyRb$ and so (3) is satisfied.

(3) $\Rightarrow$ (2) is obvious; so, the conditions (2), (3) and (4) are equivalent. Together, they imply (1): given an MST A, we obtain a hierarchy H as follows: $X \in H$: $\emptyset \notin H$; for all $x \in X$, $\{x\} \in H$; for all $x \in X$, $a \in A$, $h_{x,a} = \{z \in X\colon xzRa\} \in H$. Assume that two sets $h_{x,a}$ and $h_{y,x'}$ have a common element z. For sake of brevity, we let the reader verify that, by (2), two adjacent pairs, like xz and yz, are always comparable for R. If, say, $xzRyz$, then, for every pair a'' of the chain $A(xy)$ (which is included in $A(xz) \cup A(yz)$), one has $a''R^aa'$, $x \in h_{y,a'}$, and $h_{x,a} \subseteq h_{y,a'}$. So, H is a hierarchy. Let $a(xy)$ denotes a maximum for R in the chain $A(xy)$. It is clear that, with such a definition, $xHy = h_{x,a(xy)}$ for distinct $x, y \in X$; then, $xHy \subseteq x'Hy'$ implies $h_{x,a(xy)} \subseteq h_{x,a(x'y')}$, $xyR^sa(xy)$, $a(xy)Ra(x'y')$ and $a(x'y')R^sx'y'$. Finally, $R_H \subseteq R$. □

It appears in this proof that hierarchical DPs have strong properties. For instance, if R is an hierarchical DP and A an MST for R, then, for any $e \in X^{(2)}$, R has a maximum a in the chain $A(e)$; furthermore, $aR^s e$. We can choose such a maximum, denoted, as previously, $a(e)$. If $A(e) \subseteq A(e')$ for two pairs e and e', then $a(e)Ra(e')$ and eRe': the property $P_A \subseteq R$ is always satisfied, P_A being the dissimilarity order canonically associated with A at the end of Section 2.2.

The condition (2) above is the dissimilarity preorder version of the ultrametric inequality: a dissimilarity d on X is an ultrametric if the preorder R_d is hierarchical. This leads to the following characterization of ultrametrics in terms of MSTs [16].

COROLLARY 3.2. *A dissimilarity d on X is an ultrametric iff every pair $xy \in X^{(2)}$ belongs to an* MST *for d.*

Given $d \in \mathscr{D}$ and $A \in \mathscr{A}_{M,d}$, a dissimilarity vd is obtained by setting $vd(e) = d(e)$ for $e \in A$ and $vd(e) = \max_{a \in A(e)} d(e)$ otherwise. Obviously, vd satisfies the condition of Corollary 3.2 and is an ultrametric; the mapping v corresponds to the 'single linkage' method of the classification literature. It will be considered in the following sections.

3.2. MST*s and seriation.* Let $d \in \mathscr{D}$, and $L = x_1 < x_2 < \ldots < x_i < \ldots < x_{n-1} < x_n$ a linear order on X. The dissimilarity d and the order L are said to be *compatible,* or *semi-compatible,* if they satisfy, respectively, the following condition (C), or the weaker condition (SC):

(C) $\quad 1 \leqslant i \leqslant j < k \leqslant l \leqslant n$ implies $d(x_j, x_k) \leqslant d(x_i, x_l)$,

(SC) $\quad 1 \leqslant i \leqslant j < k \leqslant n$ implies $d(x_j, x_{j+1}) \leqslant d(x_i, x_k)$.

These conditions are easily restated in terms of MSTs. Consider the chain tree $C = \{x_j x_{j+1} : j = 1, \ldots, n-1\}$. Then, d and L are semi-compatible if C is an MST for d, and compatible when, moreover, $P_C \subseteq R_d$. In the literature of seriation, a dissimilarity d is said *Robinson* when it is compatible with some linear order L. A weakening of the Robinson property, again interesting for seriation, is that d has a chain MST, that is d is semi-compatible with some linear order L. Another generalization is the existence of a spanning tree A on X such that $P_A \subseteq R_d$; the dissimilarity d will be then said *tree Robinson.*

In the general case, MSTs provide partial seriations. Let d be a dissimilarity, $A \in \mathscr{A}_{M,d}$, and $C = \{x_1x_2, x_2x_3, \ldots, x_{k-1}x_k\}$ a chain of A. The restriction of d to the set $\{x_1, \ldots, x_k\}$ and the corresponding order on C are semi-compatible. Moreover, if d is tree Robinson with regard to A, then its restriction to C is Robinson. According to a result of Batbedat [2], this is the case, for each of its MSTs, when d is a tree dissimilarity.

It is known [5] that the clusters of a hierarchy H may be represented as intervals of a (not unique) linear order L on X. The associated chain tree C is then an MST for the DP R_H, and for every DP R with $R_H \subseteq R$; moreover, $P_C \subseteq R$. Then, every ultrametric is Robinson [6], a special case of the above Batbedat result.

4. MSTS AND CLASSES OF DISSIMILARITIES

4.1. MST-*preserving transformations of dissimilarities.* A mapping from $\mathscr{D}$ into a distinguished subset of dissimilarities that preserves MSTs also preserves the hierarchical classification associated with MSTs in Section 3.1 and the partial seriations provided by the chains of MSTs mentioned in Section 3.2. This fact constitutes a motivation for the study of such mappings.

For any $q > 0$, the subset $\mathcal{M}_q$ of the lattice $\mathcal{D}$ is closed under joins, but not under meets. Then, since the lattice $\mathcal{D}$, completed with a maximum element, is a complete lattice (where any subset, finite or infinite, has a meet and a join), there is an anticlosure operator μ_q which maps any dissimilarity d into the greatest q-semimetric $\mu_q d$ lower than d (we set $\mu_1 = \mu$). For q infinite, the set $\mathcal{U}$ is also closed under joins and the corresponding anticlosure is the previously defined mapping ν [16]. For $q = 0$, the set $\mathcal{E}$ is closed under meets as well as joins, which leads to an anticlosure ε, but also to a closure operator ε' (mapping d into the least even dissimilarity upper than d).

The q-semimetric $\mu_q d$ is also a q'-semimetric for all $q' \leqslant q$. Hence, again with the pointwise order on the anticlosure operators, we have the following.

PROPOSITION 4.1. *If $q \leqslant q'$, then $\mu_{q'} \leqslant \mu_q$.*

For $q > 0$, the condition (qM) of Section 1 is equivalent to: for all distinct $x_1, x_2, \ldots, x_k \in X$, $(x_1 x_k)^q \leqslant \sum_{1 \leqslant i \leqslant k-1} (x_i x_{i+1})^q$. It follows that, given $d \in \mathcal{D}$, the dissimilarity $\mu_q d$ may be obtained (as observed implicitly in [26]) by any minimum path length algorithm applied to the graph K_X valued by d^q. The anticlosure ε (resp. the closure ε') is obtained by iterating the operation: for each pair xy such that $d(xy) = 0$ and z distinct from x and y, set $d'(xz) = d'(yz) = \min(d(xz), d(yz))$ (resp. $\max(d(xz), d(yz))$).

A mapping $f: \mathcal{D} \to \mathcal{D}$ is said *strongly* MST-*preserving* if, for every $d \in \mathcal{D}$, $\mathcal{A}_{\mathrm{M},d} \subseteq \mathcal{A}_{\mathrm{M},f(d)}$ and, for any $A \in \mathcal{A}_{\mathrm{M},d}$, the restrictions $d_{|A}$ and $f(d)_{|A}$ of d and $f(d)$ to A are equal.

PROPOSITION 4.2. *For all $q \geqslant 0$, the anticlosure μ_q is strongly* MST-*preserving.*

PROOF. Let $d \in \mathcal{D}$, $A \in \mathcal{A}_{\mathrm{M},d}$, and a pair $xy \in A$. In any sequence $x_1, \ldots, x_k \in X$ with $x_1 = x$ and $x_k = y$, there is an x_i such that $x_i x_{i+1}$ belongs to the cocycle $D_{A,xy}$, which implies that $d(xy) \leqslant d(x_i x_{i+1})$ and $(d(xy))^q \leqslant (d(x_i x_{i+1}))^q$. The result follows for $q > 0$, the case of ultrametrics being similar. For $q = 0$, a direct proof is straightforward. □

The closure ε' is not MST-preserving in general. In fact, the set $\mathcal{M}_q$ is a lattice, with a meet, denoted here as $\triangle$, which is different from the meet $\wedge$ of $\mathcal{D}$. For $d_1, d_2 \in \mathcal{M}_q$, $d_1 \triangle d_2 = \mu_q(d_1 \wedge d_2)$ is the greatest q-semimetric lower than both d_1 and d_2. The following result makes easier the determination of an MST of $d_1 \wedge d_2$, and, thus, of $d_1 \triangle d_2$, provided that an MST A_1 for d_1 and an MST A_2 for d_2 are already known.

PROPOSITION 4.3. *With the above hypotheses, let A be an* MST *of the graph $(X, A_1 \cup A_2)$, valued by the restriction $(d_1 \wedge d_2)_{|A_1 \cup A_2}$. Then, A is an* MST *for $d_1 \wedge d_2$.*

PROOF. If $e \in A_1 \cup A_2$, then $(d_1 \wedge d_2)(e) \geqslant \max\{(d_1 \wedge d_2)(a)\colon\ a \in A(e)\}$. Otherwise, $d_1(e) \geqslant \max\{(d_1(a)\colon\ a \in A_1(e)\} \geqslant \max\{(d_1 \wedge d_2)(a)\colon\ a \in A_1(e)\} \geqslant \max\{(d_1 \wedge d_2)(a)\colon\ a \in A(e)\}$. Similarly, $d_1(e) \geqslant \max\{(d_1 \wedge d_2)(a)\colon a \in A(e)\}$. The result follows. □

Now we describe another construction [22], not related to any lattice structure at a first sight, for mapping a given dissimilarity d into a tree one $t = t_A$, with preservation of one MST A of d and equal restrictions $d_{|A}$ and $t_{|A}$. According to a result recalled in Section 3.2, the dissimilarity t is tree Robinson. It is obtained as follows:

For each $x \in X$ which is not a leaf of the tree A, set $N_x = \{y \in X\colon xy \in A\}$. Compute the dissimilarity ∂_x on N_x defined by $\partial_x(yz) = d(yz) - d(xy) - d(xz)$ and determine an MST B_x on N_x for ∂_x. Let B be the union of the sets B_x for all the elements x of

X that are not leaves of A. The values of t_A are obtained by the successive use of three rules:
(i) If $e \in A \cup B$, then $t_A(e) = d(e)$.
(ii) If $e \notin A \cup B$ and $e = yz$ with $y, z \in N_x$ for some x, then $t_A(yz) = d(xy) + d(xz) + \max\{d(vw) - d(xv) - d(xw): vw \in B_x(yz)\}$.
(iii) Otherwise, the chain $A(e)$ has at least three elements. Let $x_1, \ldots, x_{k+1}$ be its vertices in the corresponding order; then $t_A(yz) = \sum_{1 \leqslant i \leqslant k-1} d(x_i x_{i+2}) - \sum_{2 \leqslant i \leqslant k-1} d(x_i x_{i+1})$.

This algorithm is $O(n^2)$; three remarks must be made: if d has several MSTs, then t_A depends on the choice of A; there exist tree dissimilarities d with an MST A for d such that $t_A \neq d$ (fortunately, in this case, A is not unique and it is always possible to find another MST A' with $t_{A'} = d$); in the general case, one has neither $t_A \leqslant d$ nor $d \leqslant t_A$.

4.2. *Lattices of dissimilarities with fixed* MSTs. Let A be a spanning tree on X and $\mathscr{D}^A$ and $\mathscr{R}^A$ respectively be the sets of all the dissimilarities that have A as an MST and that are tree Robinson with regard to A. In other terms, $\mathscr{D}^A = \{d \in \mathscr{D}: \Delta_A \subseteq R_d\}$ and $\mathscr{R}^A = \{d \in \mathscr{D}: P_A \subseteq R_d\}$. In particular, if A is the chain-tree corresponding to a given linear order on X, $\mathscr{D}^A$ and $\mathscr{R}^A$ are, respectively, the sets of all the dissimilarities semi-compatible and compatible with this order.

Consider $d, d' \in \mathscr{D}^A$, and two pairs $a \in A$ and $b \in X^{(2)} - A$ such that $a \Delta_A b$; then, $d(a) \leqslant d(b)$ and $d'(a) \leqslant d'(b)$ imply $\min(d(a), d'(a)) \leqslant \min(d(b), d'(b))$ and $\max(d(a), d'(a)) \leqslant \max(d(b), d'(b))$; so, both $d \vee d'$ and $d \wedge d'$ belong to $\mathscr{D}^A$. Consider $d, d' \in \mathscr{R}^A$, and two pairs $e, e' \in X^{(2)}$ such that $A(e) \subseteq A(e')$; then, $d(e) \leqslant d(e')$ and $d'(e) \leqslant d'(e')$ imply $\min(d(e), d'(e)) \leqslant \min(d(e'), d'(e'))$ and $\max(d(e), d'(e)) \leqslant \max(d(e'), d'(e'))$; so, both $d \vee d'$ and $d \wedge d'$ belong to $\mathscr{R}^A$. We then have the following.

PROPOSITION 4.4. *For any $A \in \mathscr{A}$, the sets $\mathscr{D}^A$ and $\mathscr{R}^A$ are two sublattices of $\mathscr{D}$.*

Given a spanning tree A on X, there exist two anticlosures α and ρ and two closures α' and ρ' on $\mathscr{D}$ that map any dissimilarity d on dissimilarities with A as an MST; αd and $\alpha' d$ are the least element of $\mathscr{D}^A$ higher than d and the greatest lower than d, while ρd and $\rho' d$ are the least element of $\mathscr{R}^A$ higher than d and the greatest lower than d. It is straightforward to verify that these dissimilarities may be determined as follows:
(i) For $b \in X^{(2)} - A$, $\alpha d(b) = d(b)$; for $a \in A$, $\alpha d(a) = \min_{e \in D_{A,a}} d(e)$.
(ii) For $a \in A$, $\alpha' d(a) = d(a)$; for $b \in X^{(2)} - A$, $\alpha' d(b) = \max_{e \in C_{A,b}} d(e)$.
(iii) For all $e \in X^{(2)}$ such that $A(e)$ is maximal (that is, the elements of e are two leaves of A), $\rho d(e) = d(e)$; if $\rho d(e')$ is known for all e' such that $A(e) \subset A(e')$; then, $\rho d(e) = \min\{d(e), \min_{A(e) \subset A(e')} \rho d(e')\}$.
(iv) For all $a \in A$, $\rho' d(a) = d(a)$; if $\rho' d(e')$ is known for all e' such that $A(e') \subset A(e)$, then $\rho' d(e) = \max\{d(e), \max_{A(e') \subset A(e)} \rho d(e')\}$.

For $d \in \mathscr{D}$ and a spanning tree A on X (in particular, A may be an MST for d, or an objective linear order on X), there is also an ultrametric u with A as an MST, $d \leqslant u$, and u minimal with these properties. The construction is as follows:
(i) Let $m_1 = \max_{e \in X^{(2)}} d(e)$ and $M_1 = \{e \in X^{(2)}: d(e) = m_1\}$. Choose a subset $A_1 \subseteq A$ such that $M_1 \subseteq D_1 = \bigcup \{D_{A,a}: a \in A_1\}$ and D_1 is minimal with these properties. Set $u(e) = m_1$ for all $e \in D_1$, $E_2 = X^{(2)} - D_1$, $A'_2 = A - A_1$.
(ii) Repeat the previous iteration until E_{k+1} is empty: at the step k, let $m_k = \max_{e \in E_k} d(e)$ and $M_k = \{e \in E_k: d(e) = m\}$. Choose $A_k \subseteq A'_k$ such that $M_k \subseteq D_k = (\bigcup \{D_{A,a}: a \in A_1\}) \cap E_k$ and D_k is minimal with these properties. Set $u(e) = m_k$ for all $e \in D_k$, $E_{k+1} = E_k - D_k$, $A'_{k+1} = A'_k - A_k$.

The arguments are similar to those in [18]. If d has no ties (that is, no distinct pairs e and e' with $d(e) = d(e')$), the ultrametric u is minimal with the property $u \geqslant d$ alone; that is, there is no ultrametric $u' \in \mathcal{U}$ with $d \leqslant u' < u$.

5. MSTs in Lattices of Ultrametrics

In this section, we consider the particular cases of the anticlosure ν and the lattice $\mathcal{U}$. It was observed previously that the ultrametric νd is entirely defined by the restriction of d to one of its MSTs A. For all $e \in X^{(2)}$, $\nu d(e) = \max\{d(a): a \in A(e)\}$. The order on $\mathcal{U}$ may be also expressed in terms of MSTs [16]:

Proposition 5.1. *For $d, d' \in \mathcal{D}$, the following properties satisfy* (1) $\Rightarrow$ (2) $\Rightarrow$ (3):
(1) *There exists an* MST A *for d such that $d'(a) \leqslant d(a)$ for all $a \in A$.*
(2) $\nu d' \leqslant \nu d$.
(3) *For any* MST A' *for d', $d'(a) \leqslant d(a)$ for all $a \in A'$.*

Proof. If (1) is satisfied, then, for all $e \in X^{(2)}$, $\nu d'(e) \leqslant \max_{a \in A(e)} \nu d'(a) \leqslant \max_{a \in A(e)} d'(a) \leqslant \max_{a \in A(e)} d(a) = \nu d(e)$; the first inequality is due to Condition (3) of Theorem 3.1. Therefore, (1) implies (2). In turn, if (2) is satisfied, then, for all MST A' for d', and $a' \in A'$, $d'(a') = \nu d'(a') \leqslant \nu d(a') \leqslant d(a')$. □

Corollary 5.2. *For $u, u' \in \mathcal{U}$, if there exists an* MST A *for u such that $u(a) = u'(a)$ for all $a \in A$, then $u' \leqslant u$.*

For an integer $p \geqslant 1$, let $\mathcal{U}_p$ be the finite sublattice of all the ultrametrics on X with values in $\{0, 1, \ldots, p\}$. The covering relation in $\mathcal{U}_p$ is denoted as $\prec$. For $u \in \mathcal{U}$ and $A \in \mathcal{A}_{\mathrm{M},u}$, set $r(u) = \sum_{a \in A} u(a)$ (according to the ordinal definition of MSTs, this number does not depend on some choice of A). The covering pairs of elements of $\mathcal{U}_p$ satisfy the following property:

Proposition 5.3. *Let $u, u' \in \mathcal{U}_p$ with $u' \prec u$, $a_0 \in X^{(2)}$ such that $u'(a_0) < u(a_0)$, and a spanning tree $A \in \mathcal{A}_{\mathrm{M},u}$ such that $a_0 \in A$. Then, $A \in \mathcal{A}_{\mathrm{M},u'}$ and $r(u) = r(u') + 1$.*

Proof. By Corollary 3.2, the MST A exists. Define two ultrametrics u_1 and u_2 by: $A \in \mathcal{A}_{\mathrm{M},m_1}$, $u_1(a) = u(a)$ for $a \in A - \{a_0\}$ and $u_1(a_0) = u(a_0) + 1$; $A \in \mathcal{A}_{\mathrm{M},u_2}$ and $u_2(a) = u(a)$ for all $a \in A$. The ultrametrics u_1 and u_2 are well defined with, from Corollary 5.2, the inequalities $u' \leqslant u_2 \leqslant u_1 < u$. Then $u' \prec u$ implies that $u' = u_2 = u_1$. □

Theorem 5.4. *Let $u, u' \in \mathcal{U}_p$. Then u covers u' iff there exists a spanning tree $A \in \mathcal{A}_{\mathrm{M},u} \cap \mathcal{A}_{\mathrm{M},u'}$ such that $u_{|A}$ covers $u'_{|A}$ in the ordered set $\{0, \ldots, p\}^A$.*

Proof. The existence of the spanning tree A comes from Proposition 5.3. For the converse, let $A \in \mathcal{A}_{\mathrm{M},u} \cap \mathcal{A}_{\mathrm{M},u'}$ such that $u_{|A}$ covers $u'_{|A}$: $u(a) = u'(a)$ for any $a \in A$, except for a unique element a_0 such that $u(a_0) = u'(a_0) + 1$. Then, $u > u'$; assume that there exists $u_1 \in \mathcal{U}_p$ such that $u > u_1 > u'$. Since $u_{1|A} = u_{|A}$ or $u_{1|A} = u'_{|A}$, A cannot be an MST for u_1. With $A \notin \mathcal{A}_{\mathrm{M},u_1}$ and $u_{1|A} = u'_{|A}$, one has, by Corollary 5.2, $u_1 \leqslant u'$, a contradiction. If $A \notin \mathcal{A}_{\mathrm{M},u_1}$ and $u_{1|A} = u_{|A}$, there exist $x, y \in X$ such that $u_1(xy) < \max\{u_1(a): a \in A(xy)\} = \max\{u(a): a \in A(xy)\}$. Two cases may occur.

Case 1: $a_0 \notin A(xy)$, or $\max\{u(a): a \in A(xy)\} \neq u(a_0)$. Then, $u_1(xy) < u(xy) = u'(xy)$, a contradiction with $u_1 > u'$.

Case 2: $a_0 \in A(xy)$, and $\max\{u(a): a \in A(xy)\} = u(a_0)$. Then, either (2a) $u(a) < u(a_0)$ for any $a \in A(xy) - \{a_0\}$ and u, having a unique maximum on the cycle $C_{A,xy}$ in a_0, cannot be an ultrametric; or (2b) there exists $e \in A(xy)$ such that $u(e) = u(a_0)$. Then, $u'(xy) = \max\{u'(a): a \in A(xy)\} = u'(e) > u_1(xy)$, again a contradiction with $u_1 > u'$. Finally, u_1 cannot exist and u covers u'. □

THEOREM 5.5. *The lattice $\mathcal{U}_p$ is upper semimodular.*

PROOF. We have to show that, for all distinct u, u_1, $u_2 \in \mathcal{U}_p$, u_1, $u_2 \prec u$ imply $u_1 \triangle u_2 \prec u_1$, u_2. According to the previous hypotheses, there exist a_1, $a_2 \in X^{(2)}$ such that $u_1(a_1) = u(a_1) - 1$ and $u_2(a_2) = u(a_2) - 1$. Let $A \in \mathcal{A}_{\mathrm{M},u}$ such that $a_1 \in A$ and, thus, $A \in \mathcal{A}_{\mathrm{M},u_1}$ and $u_1(a) = u(a)$ for all $a \in A - \{a_1\}$. If $a_2 \in A$, then $A \in \mathcal{A}_{\mathrm{M},u_2}$ and $u_2(a) = u(a)$ for all $a \in A - \{a_2\}$. The ultrametric u' defined by $A \in \mathcal{A}_{\mathrm{M},u'}$, $u'(a) = u(a)$ for $a \in A - \{a_1, a_2\}$, $u'(a_1) = u(a_1) - 1$ and $u'(a_2) = u(a_2) - 1$ is covered by both u_1 and u_2 and so is equal to $u_1 \triangle u_2$. Otherwise, a_2 is a maximum of u in C_{A,a_2}. If there exist $a' \in C_{A,a_2}$, with $a' \neq a_1$, a_2 such that $u(a') = u(a_2)$, then the spanning tree $A_{a'a_2}$ is an MST for u containing a_1 and a_2, and the proof is as above.

Otherwise, $a_1 \in C_{A,a_2}$, $u(a_1) = u(a_2)$ and a_1 and a_2 are all the maxima of u in C_{A,a_2}. Then, the spanning tree $A_{a_1a_2}$ is an MST for u and also, by Proposition 5.3, for u_2. Furthermore, $u_1(a_2) = \max\{u_1(a): a \in A(a_2)\} = u(a_1) - 1 = u_1(a_1)$. So, A' is an MST for u_1 and, since $u_2(a_2) = u(a_2) - 1 = u(a_1) - 1 = u_1(a_2)$, we obtain $u_{1|A'} = u_{2|A'}$, which implies $u_1 = u_2$, a contradiction. □

By Theorem 5.4, the function r defined above as the length of MSTs is a rank function on the lattice $\mathcal{U}_p$. Since a finite lattice is lower semimodular iff its rank function is a lower valuation on the lattice $\mathcal{U}_p$ [23], we then have:

$$\text{For all } u, u' \in \mathcal{U}_p,\ r(u) + r(u') \leqslant r(u \vee u') + r(u \triangle u').$$

The use of decimal approximation with preservation of the order of values makes it possible to extend this property to the function r defined on the whole lattice $\mathcal{U}$ (problem: find a simple direct proof of this result). Similarly, the following characterization of the semimodularity property is due to Dubreil-Jacotin, Lesieur and Croisot [7]; it extends the definition of semimodularity to infinite lattices:

For all u, u', $u'' \in \mathcal{U}$, $u' \vee u'' > u > u'' > u \triangle u'$ imply that there exists $v \in \mathcal{U}$

$$\text{such that } u' \vee u'' > v \geqslant u' \text{ and } u = (u \triangle v) \vee u''.$$

It is then straightforward to verify that the infinite lattice $\mathcal{U}$ is semimodular in that sense; the status of this lattice as a kind of product of two lattices, $\mathbb{R}_+$ and the partition lattice, is explained in [21]. More generally, the question arises of the algebraic properties of lattices $\mathcal{M}_q$.

As a final remark, it is worth noticing that, if a spanning tree A is a common MST for two ultrametrics u and u', then A is still an MST for $u \vee u'$ (by Proposition 4.4) and for $u \triangle u'$ (by Proposition 4.3). Furthermore, $(u \vee u')(a) = \max(u(a), u'(a))$ and $(u \triangle u')(a) = \min(u(a), u'(a))$ for all $a \in A$. Therefore a distributive sublattice $\mathcal{U}^A$ of $\mathcal{U}$, isomorphic to $(\mathbb{R}_+)^{n-1}$, is associated to any tree A on X (as noted in [1]; $\mathcal{U}^A$ is the lattice of all the ultrametrics with A as an MST. In such a case of distributivity, aggregation problems become easier [20, 24].

References

1. J. P. Barthélemy, B. Leclerc and B. Monjardet, on the use of ordered sets in problems of comparison and consensus of classification, *J. Classification* **3** (1986), 187–224.
2. A. Batbedat, *Les Approches Pyramidales dans la Classification Arborée,* Masson, Paris, 1990.
3. G. Birkhoff, *Lattice Theory,* 3rd edn, Am. Math. Soc., Providence, RI, (1967).
4. O. Boruvka, O jistému problému minimálním (On a minimal problem), *Prace Moravské Prirodovedecké Spolecnosti v Brne* **3** (1926), 37–58.
5. G. Brossier, Représentation ordonnée des classifications hiérarchiques, *Stat. Anal. Données,* **2** (1980), 31–44.
6. E. Diday, Croisements, ordres et ultramétriques, *Math. Sci. Hum.,* **83** (1983), 31–54.
7. M. L. Dubreil-Jacotin, L. Lesieur and P. Croisot (1953), *Leçons sur la Théorie des Treillis,* Cahiers Scientifiques 21, Gauthier-Villars, Paris.
8. C. Flament and B. Leclerc, Arbres minimaux d'un graphe préordonné, *Discr. Math.,* **46** (1983), 159–171.
9. K. Florek, J. Łukaszewicz, H. Perkal, H. Steinhaus and S. Zubrzycki, Sur la liaison et la division des points d'un ensemble fini, *Colloq. Math.* **20** (1951), 282–285.
10. D. Gale. Optimal assignments in an ordered set: an application of matroid theory, *J. Combin. Theory,* **4,** (1968), 176–180.
11. J. C. Gower and G. J. S. Ross, Minimum spanning tree and single linkage cluster analysis, *Appl. Stat.* **18** (1969), 54–64.
12. R. L. Graham and P. Hell. On the history of the minimum spanning tree problem, *Ann. Hist. Comput.* **7**(1) (1985), 43–57.
13. A. Guénoche, Spanning trees and average linkage clustering, in: *New Approaches in Classification and Data Analysis,* E. Diday *et al.* (eds), Studies in Classification, Data Analysis, and Knowledge Organization, Springer-Verlag, Berlin, 1996, pp. 119–127.
14. A. Guénoche, P. Hansen and B. Jaumard, Efficient algorithms for divisive hierarchical clustering with the diameter criterion, *J. Classification,* **8** (1991), 5–44.
15. P. Hansen and B. Jaumard. Minimum sum of diameters clustering, *J. Classification,* **4** (1987), 215–226.
16. B. Leclerc, Description combinatoire des ultramétriques, *Math. Sci. Hum.,* **73** (1981), 5–37.
17. B. Leclerc, Les hiérarchies de parties et leur demi-treillis, *Math. Sci. Hum.,* **89** (1985), 5–34.
18. B. Leclerc, Caractérisation, dénombrement et construction des ultramétriques supérieures minimales, *Stat. Analyse Données,* **11** (2) (1986), 26–50.
19. B. Leclerc (1992), Arbres, ordres, treillis; contributions à l'analyse combinatoire des données, Thesis and R. R. CMS-P.082, C.A.M.S., Paris.
20. B. Leclerc, Medians for weight metrics in the covering graphs of semilattices, *Discr. Appl. Math.,* **49** (1994), 281–297.
21. B. Leclerc, The residuation model for the ordinal construction of dissimilarities and other valued objects, in: *Classification and Dissimilarity Analysis,* B. Van Cutsem (ed.), Lecture Notes in Statistics 93, Springer-Verlag, New York, 1994, pp. 149–172.
22. B. Leclerc, Minimum spanning trees for tree metrics: abridgements and adjustments, *J. Classification,* **12** (1995), to appear.
23. B. Monjardet, Metrics on partially ordered sets—A survey, *Discr. Math.,* **35** (1980), 173–184.
24. B. Monjardet, Arrowian characterization of latticial federation consensus functions, *Math. Soc. Sci.,* **20** (1990), 51–71.
25. R. Rado, Note on independence functions, *Proc. Lond. Math. Soc.* **7** (1957), 300–320.
26. B. Van Cutsem, Ultramétriques, distances, ϕ-distances maximum dominées par une dissimilarité donnée, *Stat. Anal. Données,* **8** (1983), 42–63.

Received 30 January 1995 and accepted in revised form 6 June 1995

Bruno Leclerc
Centre d'Analyse et de Mathématique Sociales,
École des Hautes Études en Sciences Sociales,
54bd Raspail, 75270 Paris Cedex 06, France
E-mail: *leclerc@ehess.fr*